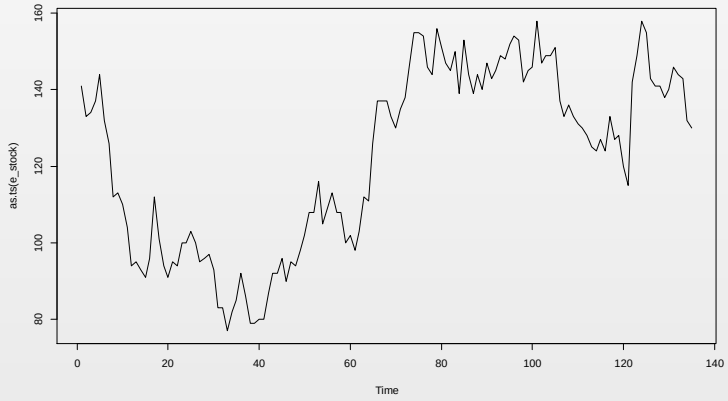
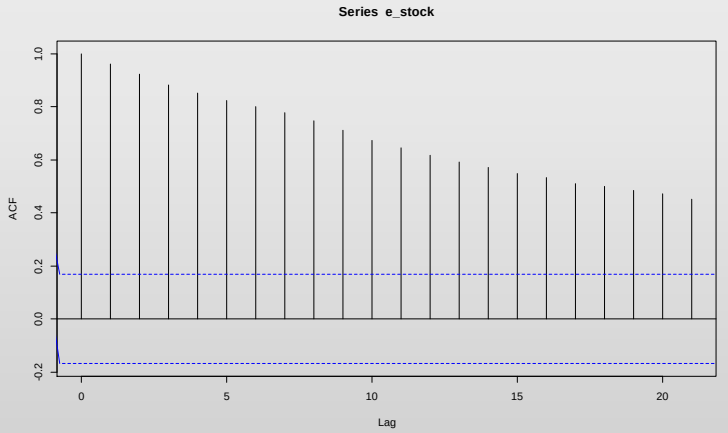
예 8-8(316쪽): 데이터 파일 elecstock.txt

```
> e_stock <- scan("D:/Data/elecstock.txt")
Read 135 items
```



시계열 그림

비정상 시계열
차분 필요



ACF

- 단위근 검정에 의한 차분 필요성 결정

Augmented Dickey-Fuller Unit Root Test(ADF test)의 실시

```
> library(tseries)
> adf.test(e_stock)

        Augmented Dickey-Fuller Test

data:  e_stock
Dickey-Fuller = -2.6174, Lag order = 5, p-value = 0.3197
alternative hypothesis: stationary
```

- 귀무가설은 비정상 시계열(즉, 단위근을 갖고 있음)
  → 귀무가설을 기각할 수 없음. 차분 필요.

- 교재 317쪽 표 8-6의 SAS 결과보다는 훨씬 간략한 결과

- 함수 ndiffs( )에 의한 최적 차분 차수 결정

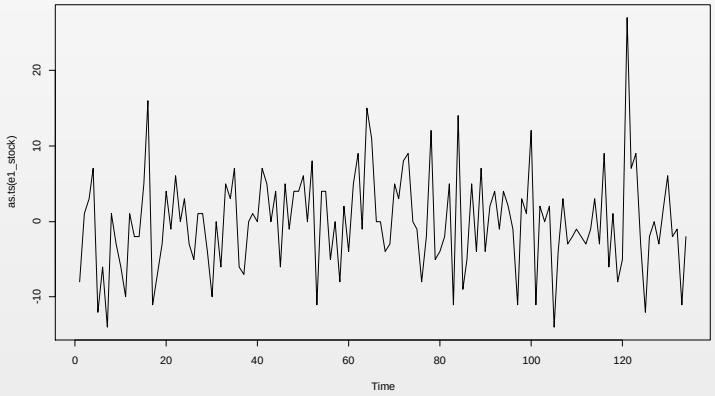
```
> library(forecast)
> ndiffs(e_stock)
[1] 1
```

  - 함수 ndiffs( ): 단위근 검정을 실시하여 정상성을 확보하는데 필요한 차분의 차수를 결정

  - 차분의 필요성 및 차분 차수를 결정하는 것이라면 굳이 함수 adf.test( )를 사용해야 할 이유 없음

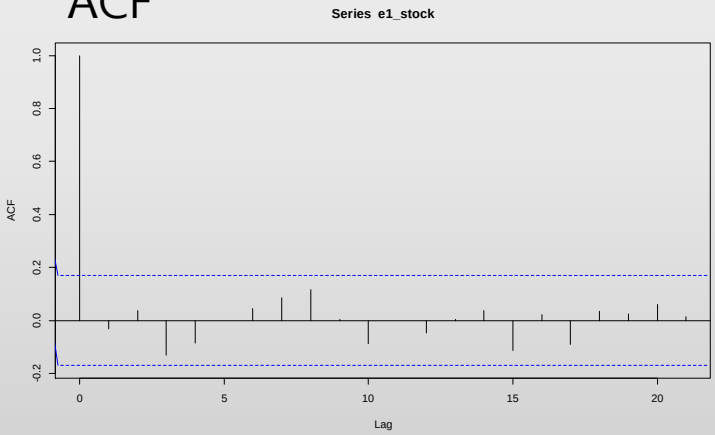
- 1차 차분 자료에 대한 모형 인식

```
> e1_stock <- diff(e_stock)
```

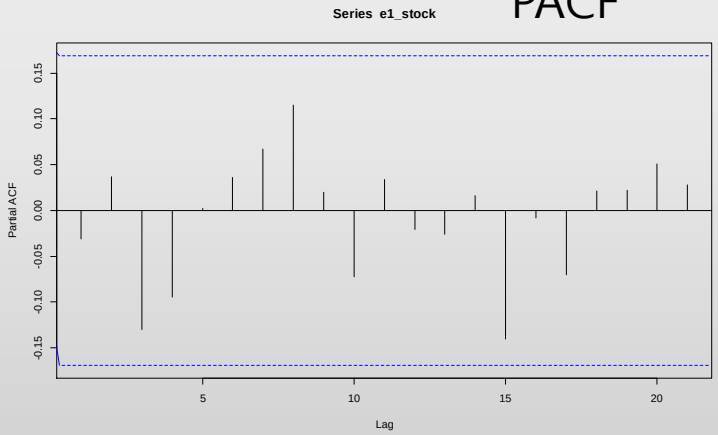


- 1차 차분된 자료 → 백색잡음과정
- 원 자료 → 확률보행과정

ACF



PACF



- 절편 포함 여부 확인
  - 차분된 자료의 경우 모형에 절편 포함 여부 결정은 필수적
  - 함수 arima( )는 d≥1인 경우 절편의 추정 및 검정은 불가능
  - 이러한 경우 패키지 forecast의 함수 Arima( ) 이용하여 검정

```
Arima( x , order=c(0,0,0) , include.mean=TRUE , include.drift=FALSE ,
          fixed=NULL )

  - x : 시계열 자료
  - order=c(0,0,0) :  ARIMA(p,d,q)의 차수 지정
  - include.mean : d=0의 자료에 대하여 모형의 평균 포함 여부
  - include.drift : d=1의 자료에 대하여 절편 포함 여부
  - fixed : 비유의적 모수 제거
```

- 원 자료에 확률보행과정 적합

```
> Arima(e_stock,order=c(0,1,0),include.drift=TRUE)
Series: e_stock
ARIMA(0,1,0) with drift

Coefficients:
         drift
      -0.0821
s.e.   0.5615

sigma^2 estimated as 42.57:  log likelihood=-440.97
AIC=885.93   AICc=886.02   BIC=891.73
```

```
> (fit <- Arima(e_stock,order=c(0,1,0)))
Series: e_stock
ARIMA(0,1,0)

sigma^2 estimated as 42.26:  log likelihood=-440.98
AIC=883.95   AICc=883.98   BIC=886.85
```

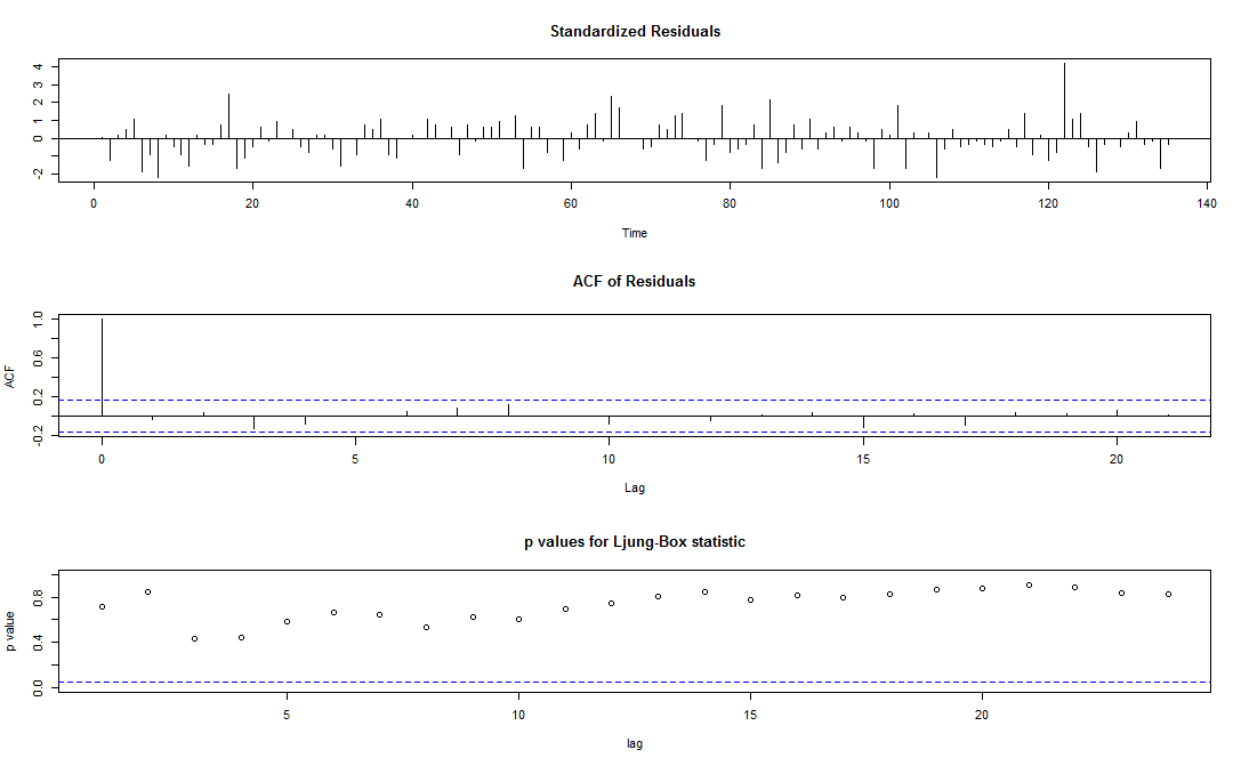
모형식

$$Z_t = Z_{t-1} + \varepsilon_t$$

$$\hat{\sigma}^2 = 42.26$$

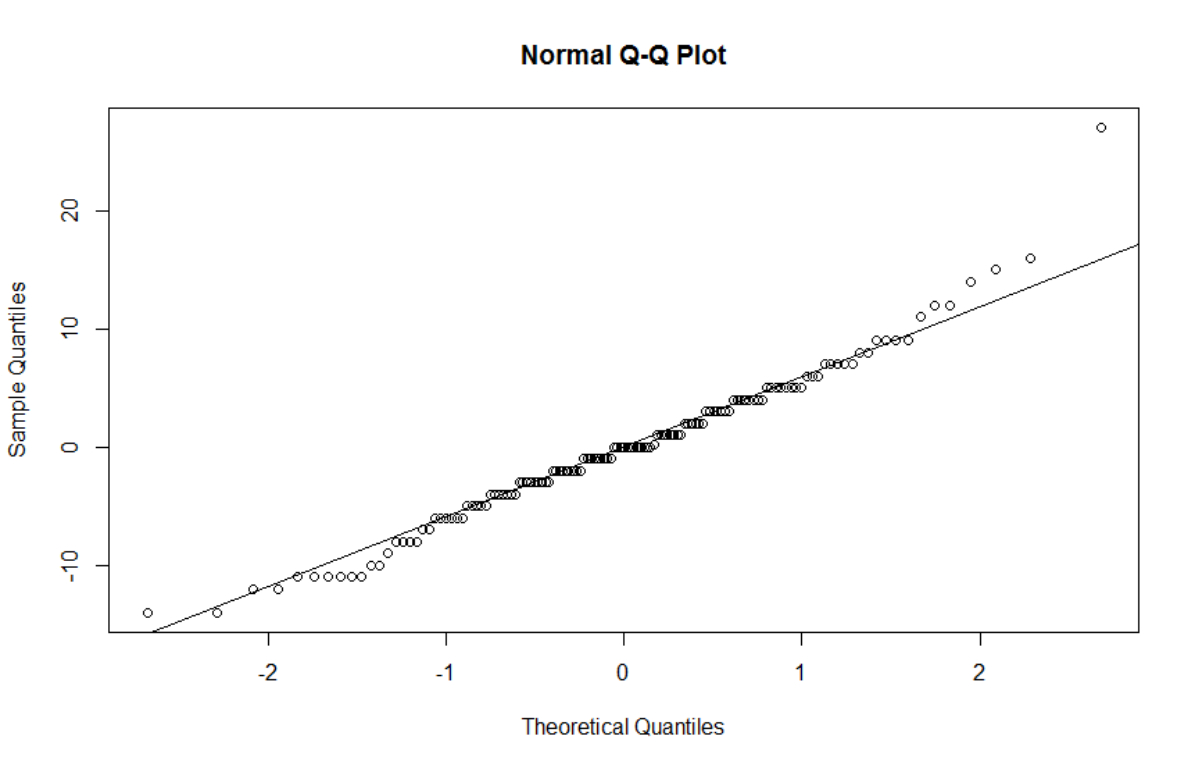
- 모형 검진

```
> tsdiag(fit, gof.lag=24)
```



- 정규분포 확인

```
> qqnorm(fit$resid) ; qqline(fit$resid)
```



- ARIMA(0,1,0) 모형에 과대적합

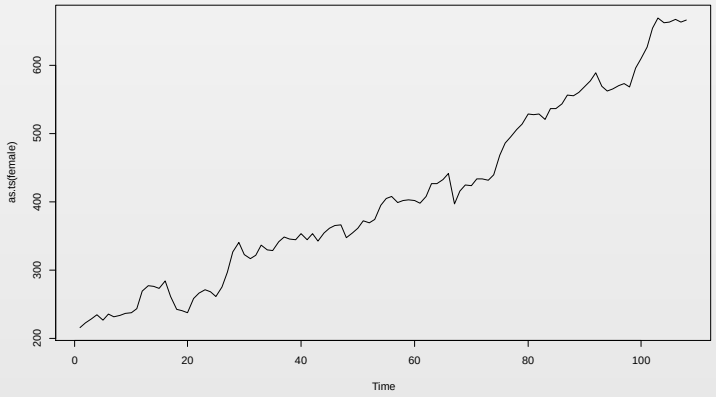
```
> confint(Arima(e_stock,order=c(1,1,0)))
          2.5 %    97.5 %
ar1 -0.2010195 0.1382239

> confint(Arima(e_stock,order=c(0,1,1)))
          2.5 %    97.5 %
ma1 -0.1950423 0.1359215
```

추가된 모수 모두 비유의적 → 절편 없는 ARIMA(0,1,0) 모형 확정

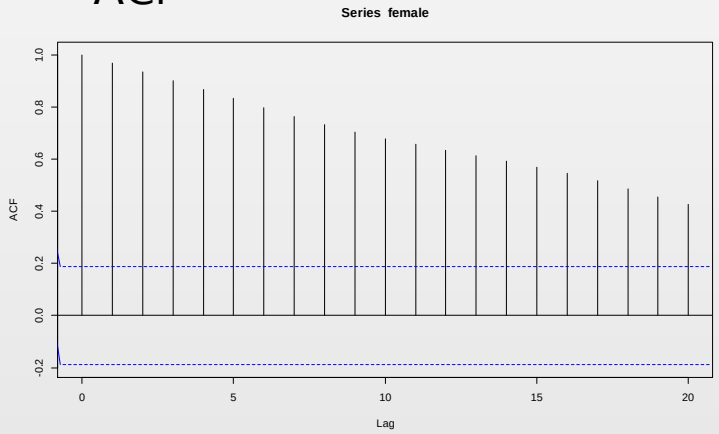
예제 8-9(319쪽): 데이터 파일 female.txt

```
> female <- scan("D:/Data/female.txt")
Read 108 items
```

시계열 그림



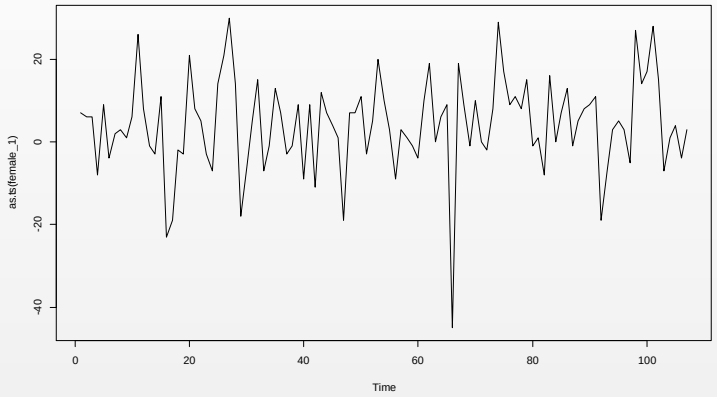
ACF



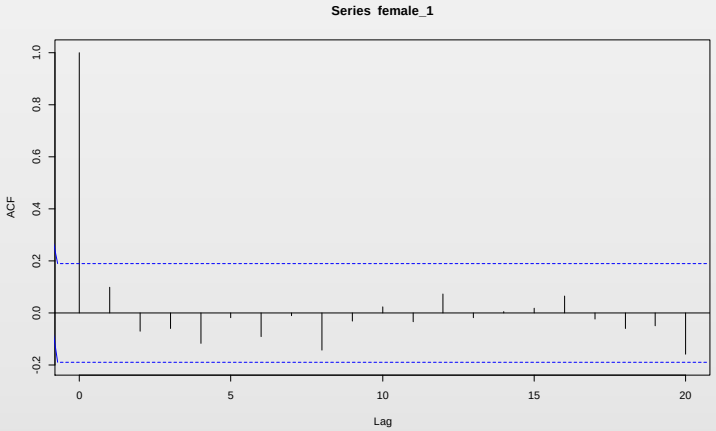
비정상 시계열 → 차분 필요

```
> ndiffs(female)
[1] 1
```

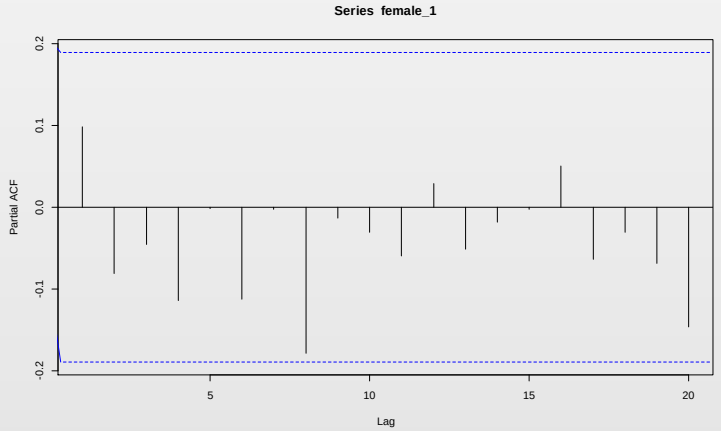
- 1차 차분된 자료: 모형 인식



백색 잡음과정



ACF



PACF

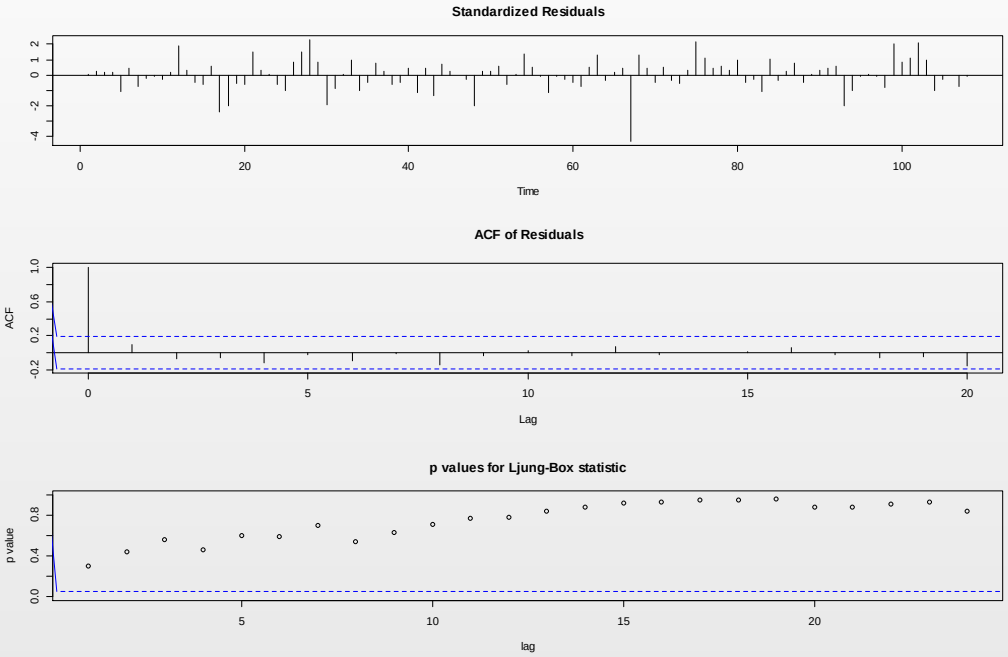
- 확률보행과정 적합

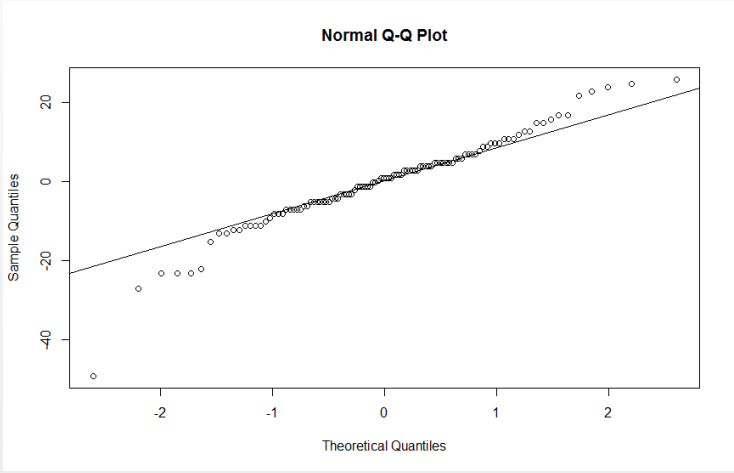
```
> fit <- Arima(female,order=c(0,1,0),include.drift=TRUE)
> fit
Series: female
ARIMA(0,1,0) with drift

Coefficients:
      drift
      4.215
s.e.  1.093

sigma^2 estimated as 129:  log likelihood=-411.34
AIC=826.68   AICc=826.8   BIC=832.03
```

모형식 $Z_t = 4.215 + Z_{t-1} + \varepsilon_t$

$\hat{\sigma}^2 = 129$

- 모형 검진


Standardized Residuals
Time
ACF of Residuals
ACF
Lag
p values for Ljung-Box statistic
p value
lag


Normal Q-Q Plot
Sample Quantiles
Theoretical Quantiles

- 과대적합

```
> confint(Arima(female,order=c(1,1,0),include.drift = TRUE))
              2.5 %    97.5 %
ar1    -0.09026514 0.2851598
drift   1.85683163 6.5762383

> confint(Arima(female,order=c(0,1,1),include.drift = TRUE))
              2.5 %    97.5 %
ma1    -0.08681113 0.3124909
drift   1.84910777 6.5857555
```

추가된 모수 모두 비유의적 → 절편 있는 ARIMA(0,1,0) 모형 확정

연습문제

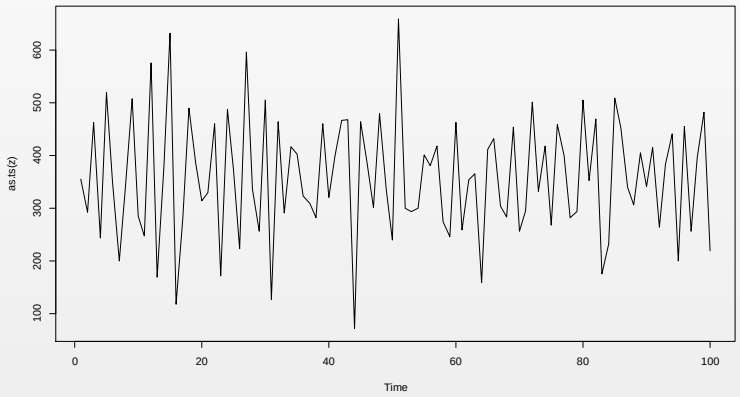
8.2 다음의 모의 시계열 자료를 대상으로 적절한 ARIMA 모형을 식별, 추정한 후 모형진단을 통해 가장 적합한 모형을 구하라.

자료:

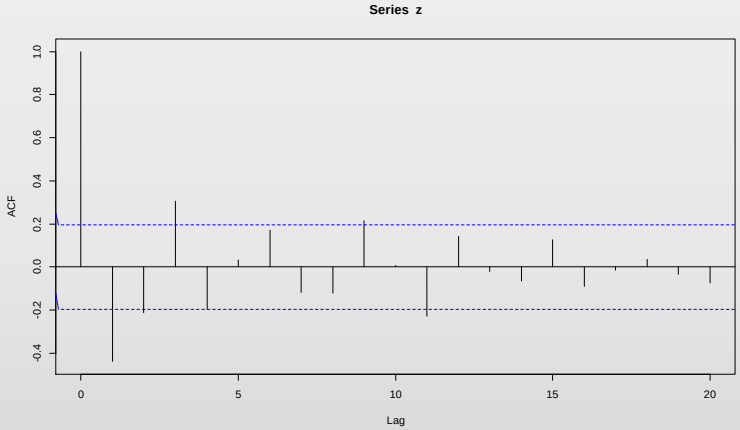
ex8_2a.txt

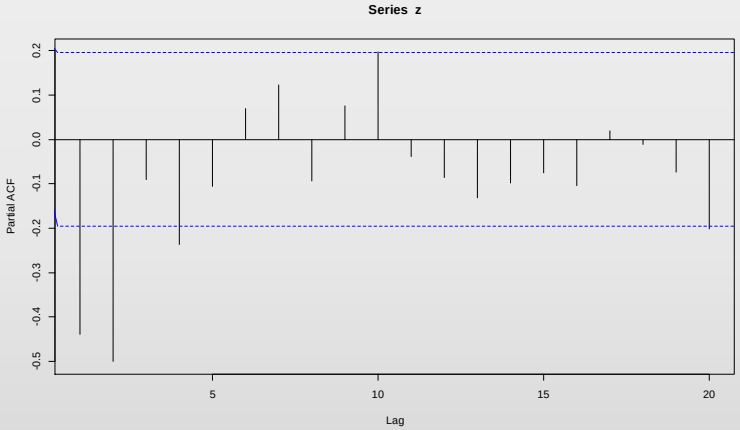
ex8_2b.txt

ex8_2a.txt



AR(4), MA(3), ARMA





- 인식된 모형의 AIC, BIC 비교

```
> arima(z,order=c(0,0,3))$aic
[1] 1183.683
> arima(z,order=c(4,0,0))$aic
[1] 1184.899
> arima(z,order=c(1,0,1))$aic
[1] 1191.028
> arima(z,order=c(1,0,2))$aic
[1] 1188.176
> arima(z,order=c(2,0,1))$aic
[1] 1186.561
> arima(z,order=c(2,0,2))$aic
[1] 1186.222
```

```
> library(forecast)
> Arima(z,order=c(0,0,3))$bic
[1] 1196.709
> Arima(z,order=c(4,0,0))$bic
[1] 1200.53
> Arima(z,order=c(1,0,1))$bic
[1] 1201.449
> Arima(z,order=c(1,0,2))$bic
[1] 1201.202
> Arima(z,order=c(2,0,1))$bic
[1] 1199.587
> Arima(z,order=c(2,0,2))$bic
[1] 1201.853
```

- MA(3) 모형의 AIC & BIC가 최소
- AR(4) & ARMA(2,1) 고려 대상

함수 arima( )에서는
BIC 출력이 안 됨

1) MA(3) 모형의 적합

```
> fit1 <- arima(z,order=c(0,0,3))
> confint(fit1)
                  2.5 %       97.5 %
ma1          -0.9385840  -0.55590008
ma2          -0.4610481   0.04668583
ma3           0.1585552   0.57494836
intercept 351.6659267 365.44390137
```

- 비유의적인 모수 존재 → 제거
- 함수 arima( )에서 옵션 fixed 사용

옵션 fixed의 사용법

- 예를 들어 절편이 있는 ARMA(2,2)에서 AR 모수 중 ϕ_1이 비유의적이라 제외하고자 한다면 다음과 같이 지정

arima(z, order=c(2,0,2),fixed=c(0, NA, NA, NA, NA))

AR모수 MA모수 절편

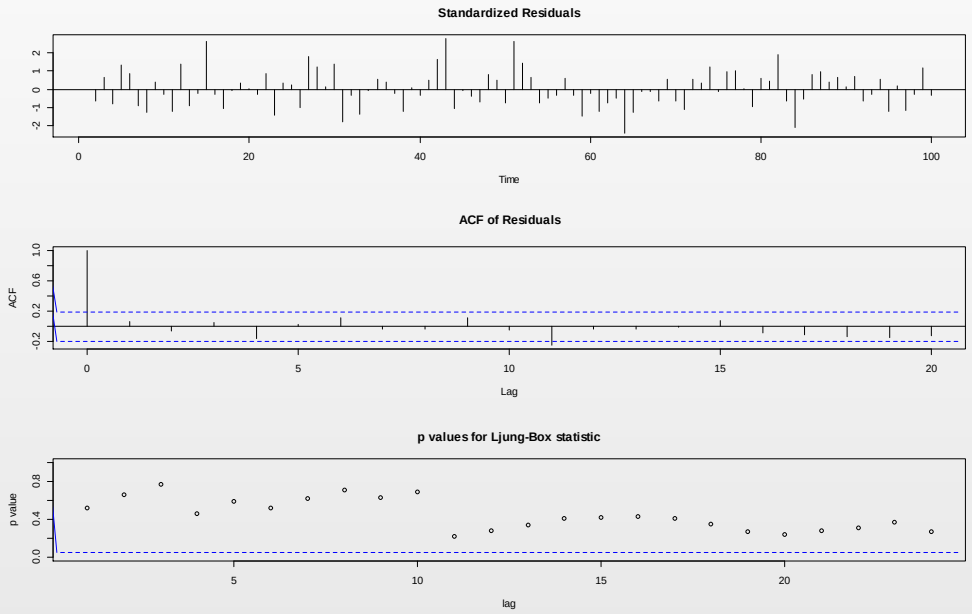
- 비유의적 모수 제거 후 모형 적합

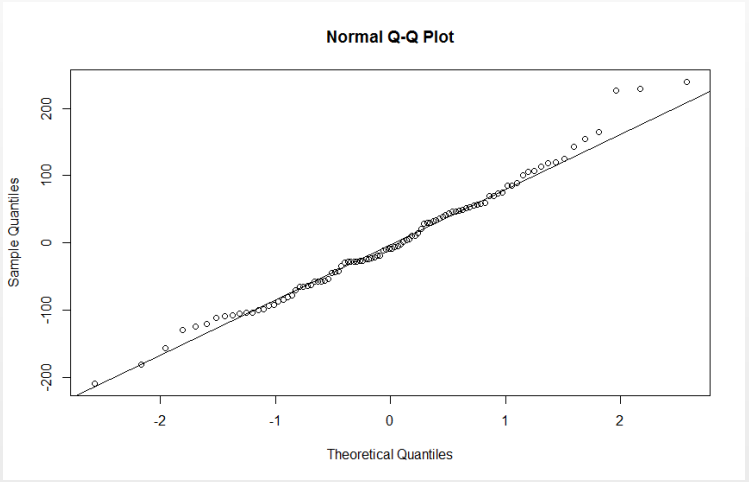
```
> fit1_2 <- arima(z,order=c(0,0,3),fixed=c(NA,0,NA,NA))
> fit1_2

Call:
arima(x = z, order = c(0, 0, 3), fixed = c(NA, 0, NA, NA))

Coefficients:
          ma1  ma2     ma3  intercept
      -0.8534    0  0.2458   358.5885
s.e.   0.0765    0  0.0777     3.3863

sigma^2 estimated as 7396:  log likelihood = -587.88,  aic = 1183.75
```

모형식

$$Z_t - 358.58 = \varepsilon_t - 0.8534\varepsilon_{t-1} + 0.2458\varepsilon_{t-3}$$

- 모형 검진


Standardized Residuals
Time
ACF of Residuals
ACF
Lag
p values for Ljung-Box statistic
p value
lag


Normal Q-Q Plot
Sample Quantiles
Theoretical Quantiles

- MA(3) 모형의 과대적합

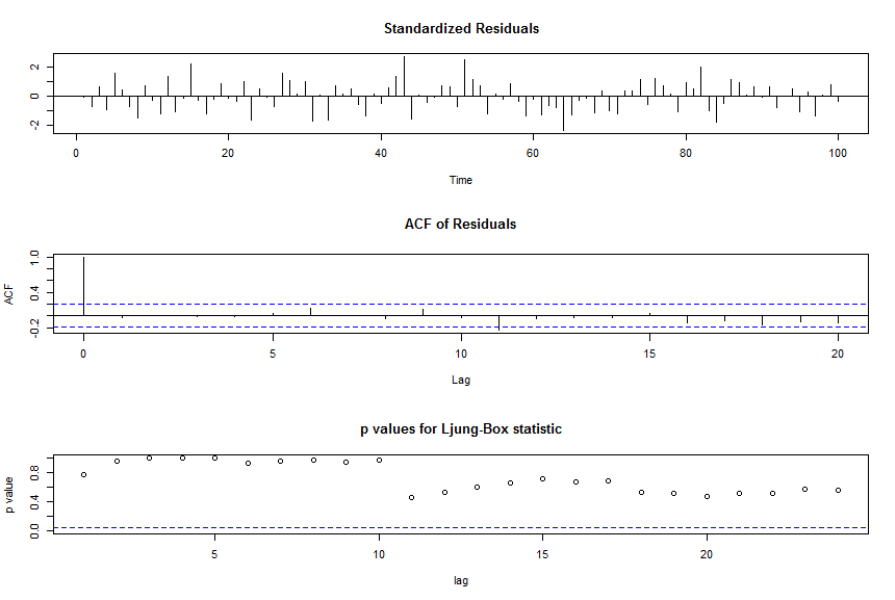
```
> confint(arima(z,order=c(0,0,4),fixed=c(NA,0,NA,NA,NA)))
                  2.5 %      97.5 %
ma1        -1.01163161  -0.6980373
ma2                 NA          NA
ma3         0.07091272   0.4544947
ma4        -0.22464932   0.1669034
intercept 352.18772417 365.0280467

> confint(arima(z,order=c(1,0,3),fixed=c(NA,NA,0,NA,NA)))
                 2.5 %      97.5 %
ar1        -0.1329726   0.4450242
ma1        -1.1229593  -0.7466816
ma2                NA          NA
ma3         0.1424891   0.4513128
intercept 351.3253347 365.7072609
```

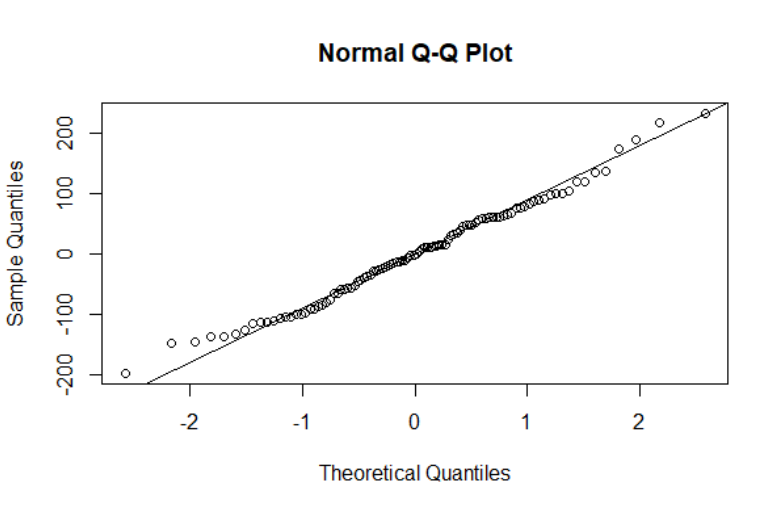
추가된 모수 모두 비유의적 → MA(3) 모형 최종모형으로 선택

2) AR(4) 모형의 적합

```
> fit2 <- arima(z, order=c(4,0,0))
> fit2

Coefficients:
          ar1      ar2      ar3      ar4    intercept
      -0.7229  -0.6888  -0.2497  -0.2375     358.6815
s.e.   0.0970   0.1182   0.1179   0.0968       2.9647

sigma^2 estimated as 7191:  log likelihood = -586.45,  aic =
1184.9

> confint(fit2)
                 2.5 %          97.5 %
ar1          -0.9130224    -0.53274787
ar2          -0.9204464    -0.45709193
ar3          -0.4808213    -0.01858373
ar4          -0.4271988    -0.04773790
intercept   352.8707181   364.49222589
```

- 모형 검진


Standardized Residuals
Time
ACF of Residuals
ACF
Lag
p values for Ljung-Box statistic
p value
lag


Normal Q-Q Plot
Sample Quantiles
Theoretical Quantiles

- AR(4) 모형의 과대적합

```
> confint(arima(z,order=c(5,0,0)))
                2.5 %       97.5 %
ar1        -0.9417832  -0.55132560
ar2        -0.9495981  -0.47736526
ar3        -0.5856012  -0.04921727
ar4        -0.5393472  -0.07155415
ar5        -0.2936309   0.10007168
intercept 353.3478236 363.91608834

> confint(arima(z,order=c(4,0,1)))
                2.5 %       97.5 %
ar1        -1.0716495   0.01162329
ar2        -0.9835229  -0.12725061
ar3        -0.5133456   0.22188390
ar4        -0.4276617  -0.02285273
ma1        -0.7452024   0.33249111
intercept 353.2156852 364.09098895
```

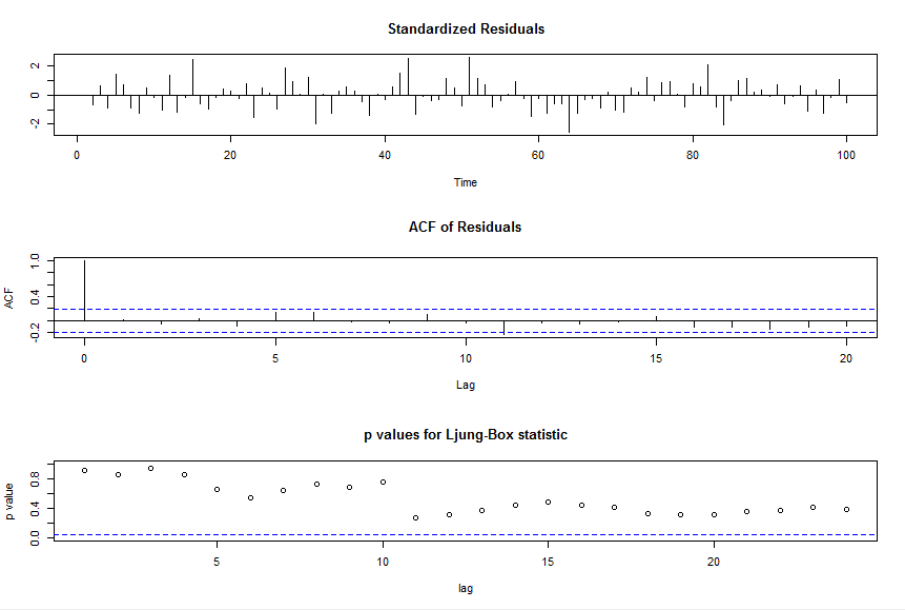
- 추가된 모수: 비유의적

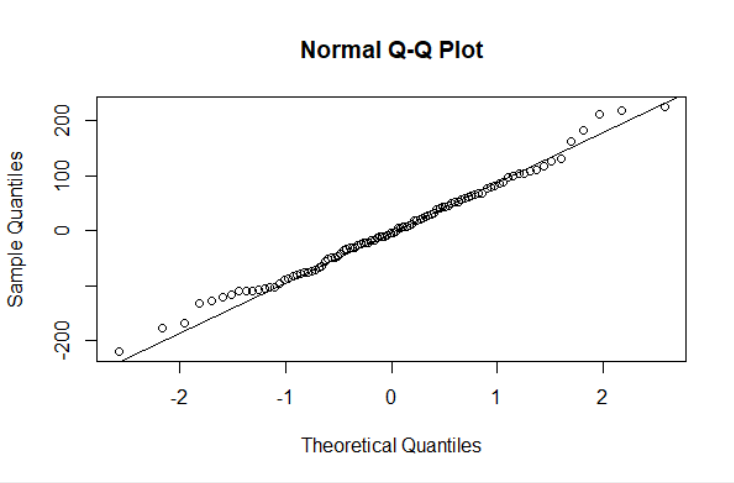
- AR(4) 최종모형으로 사용 가능

3) ARMA(2,1) 모형의 적합

```
> fit3 <- arima(z,order=c(2,0,1))
> fit3

Call:
arima(x = z, order = c(2, 0, 1))

Coefficients:
          ar1      ar2      ma1  intercept
      -0.3360  -0.3441  -0.4362   358.6120
s.e.   0.1644   0.1258   0.1686     2.9417

sigma^2 estimated as 7470:  log likelihood = -588.28,  aic =
1186.56
```

- 모형 검진


Standardized Residuals
Time
ACF of Residuals
ACF
Lag
p values for Ljung-Box statistic
p value
lag


Normal Q-Q Plot
Sample Quantiles
Theoretical Quantiles

- ARMA(2,1) 모형의 과대적합

```
> confint(arima(z,order=c(3,0,1)))
                 2.5 %      97.5 %
ar1        -0.6742690   0.3352320
ar2        -0.6058290   0.1532975
ar3        -0.2066342   0.4707361
ma1        -1.0406437  -0.1350844
intercept 353.0273959 364.2213366

> confint(arima(z,order=c(2,0,2)))
                 2.5 %        97.5 %
ar1        -1.1522263  -0.347633608
ar2        -0.5901331  -0.113922790
ma1        -0.3952026   0.431793493
ma2        -0.6803903   0.001852668
intercept 353.1589311 364.158805617
```

ARMA(3,1):
무의미한 모형

ARMA(2,2):
추가된 모수 비유의적

ARMA(2,1):
최종 모형으로 사용 가능

4) 모형의 비교

```
> fit1_2
Coefficients:
          ma1  ma2     ma3  intercept
      -0.8534    0  0.2458   358.5885
s.e.   0.0765    0  0.0777     3.3863
sigma^2 estimated as 7396:  log likelihood = -587.88,  aic = 1183.75
```

MA(3)

```
> fit2
Coefficients:
          ar1     ar2     ar3     ar4  intercept
      -0.7229 -0.6888 -0.2497 -0.2375   358.6815
s.e.   0.0970  0.1182  0.1179  0.0968     2.9647
sigma^2 estimated as 7191:  log likelihood = -586.45,  aic = 1184.9
```

AR(4)

```
> fit3
Coefficients:
          ar1     ar2     ma1  intercept
      -0.3360 -0.3441 -0.4362   358.6120
s.e.   0.1644  0.1258  0.1686     2.9417
sigma^2 estimated as 7470:  log likelihood = -588.28,  aic = 1186.56
```

ARMA(2,1)